Calhoun
Institutional Archive of the Naval Postgraduate School

Calhoun: The NPS Institutional Archive

DSpace Repository

Faculty and Researchers | Faculty and Researchers' Publications

1993-05

Notes on Three Kinds of Science

Hamming, Richard W.

Monterey, California: Naval Postgraduate School

http://hdl.handle.net/10945/63719

This publication is a work of the U.S. Government as defined in Title 17, United States Code, Section 101. Copyright protection is not available for this work in the United States.

Downloaded from NPS Archive: Calhoun

DUDLEY KNOX LIBRARY

Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

http://www.nps.edu/library

Notes on

THREE KINDS OF SCIENCE

R. W. Hamming, May 1993

There are many ways of classifying science. The value of a classification is in how much it illuminates the past and how fruitful it is in suggesting new approaches.

Science arose to a great extent in the attempt to find regular patterns in our experiences. The recognition of patterns is essential to the control of the future; and we use this every day when we handle things, take a step, build a bridge, etc..

In the earlier science the general belief was that Nature was regular and the small differences we saw were due to some kind of "noise" in the observing process.

Probability and statistics are relatively new tools for coping with randomness in Nature. When Maxwell published his two great papers on the molecular basis of gas behavior he launched a new way of systematically examining Nature, though earlier efforts, from the Bernoullis on, had played with the idea of a random basis for Nature. Mendel had also postulated a random element in Nature, as had many others, including gamblers.

There is a third kind of science that is slowly growing but is seldom recognized. A good example is the distribution of the prime numbers. First, we believe that the primes are what they are, and in no sense are they random. Second, Maxwell assumed a set of initial positions and velocities for these molecules, but since he could not hope to measure them he was immediately led to the idea of an ensemble of cases, and he dealt with the typical, average (over the emsemble) behavior. His contribution included looking at the distribution of velocities rather than assuming that they were all the same velocity for each molecule.

It is not reasonable to assume that the distribution of the prime numbers are a sample from some ensemble as is typical in classical statistics; they are what they are and there are no other similar cases to form an ensemble, though some statisticians cheerfull do so. The idea that you can get to the classical statistics via starting at a random place, and ending at another random place is not really sensible.

Yet we have asked the question, "To what extent can we regard this perfectly determinate, unique sequence as if it were random?" The classical formula for the cummulative distribution of the primes is an example of what can emerge. It has suggsted tings like the distribution of prime pairs, and prime triples (both left hand and right hand), though it remaind to compare them with the actual data. Similarly, we have found that the cummulative distribution of the leading digits of many mathematically generated sequences like n!, the Bernoulli numbers, the Fibonacci numbers, and even the succesive powers of a number (whose $\log_e$ is not rational) all seem to be close to the reciprocal distribution, but the primes, the k-th powers of the integers are not.

This third kind of science is similar to, but significantly distinct from, statistics that we need both to take what we can from statistics and also realize that we must redefine things to suit our new needs. Consider, example, the actual data from the primes as compared with the mathematically-derived cummulative distribution. Most of the data we have falls below the theoretical distribution but it has apparently been proven that infinitely often the data must be above. How shall we measure this deviation from the theory? Modelling on the very useful idea of variance we could adopt the formula that it is the square of the difference from the theory (not from the average) that is a measure. Since the mathematics of least squares, and in general quadratic forms, is fairly easy to do, it is likely that at first this measure may lead to derivable results. Again, shall we count the difference at every integer, or only at the primes where a change occurs? The test must be: what kinds of things can we derive, and if it turns out that L_1 or L_∞ is easier to work with than L_2, or give more interesting results, then we will have to adopt the corresponding definition. As Einstein is quoted by Heisenberg, "The theory determines what you see."

As we feel our way into this new field of looking at perfectly determinate, unique sequences of numbers as if they were random (in some sense or other, again open to new definitions if convenient) the choices we make in adapting classical statistical definitions and methods must be mainly determined by the kind of things we believe can be later accounted for by mathematical derivations. I said "new", but, as the example of the distribution of the primes shows, it already has a long history; it is just that I am claiming that there is a whole body of theory to be developed beyond the special cases already explored, and which assure us that there is something there to be found more general than the special cases.

Thus I claim that there is a third kind of science in the making, the science of looking at determinate sequences of numbers, or other kinds of events, as if they were random, and finding out in what respects, and for what kinds of measurements on them, this is reasonable or not. Can this approach also shed light on the output of a production line which is what it is at the time you look at it, (and the pretense that it is a member of an ensemble is dubious)? We can but try, and hope that from this different view point we will get new light on important problems beyond those real of theoretical interest. ~~mainly~~.

why limit it at 3?
other axes of classification
guided vs. free

	determinate	random	
process	classics	?	chaos?
distribution	"new" ~~paper~~	statistical	

guided determinate process — mathematics
guided determinate distribution — evolution
guided random process — fault mode
guided random distribution — economics

free determinate process — ~~chemistry~~ inorganic chem
free determinate distribution — gas physics
free random process — organic chem
free random distribution — classical statistics